

On the Riemann Hypothesis and the Difference Between Primes

ADRIAN W. DUDEK
 Mathematical Sciences Institute
 The Australian National University
 adrian.dudek@anu.edu.au

Abstract

We prove two results on the assumption of the Riemann hypothesis. First, we show that there exists a prime in the interval $(x - \frac{4}{\pi}\sqrt{x}\log x, x]$ for all $x \geq 2$, which improves a result of Ramaré and Saouter. We then show that the number of primes in the interval

$$(x - (c + 1)\sqrt{x}\log x, x + (c + 1)\sqrt{x}\log x]$$

is greater than \sqrt{x} for $c = 2/\pi$ and for all $x \geq 2$. This improves a theorem of Goldston.

1 Introduction

There is a rich interplay between the zeroes of the Riemann zeta-function $\zeta(s)$ and the distribution of prime numbers, best exemplified in the proof of the prime number theorem (see Ingham's classic text [3] for more details). The Riemann hypothesis, which asserts that all of the zeroes of $\zeta(s)$ are either negative even numbers or complex numbers with a real part of $1/2$, thus serves as somewhat of a holy grail of number theory.

On the assumption of the Riemann hypothesis, von Koch [10] proved that there exists a constant k such that the interval $(x - k\sqrt{x}\log^2 x, x)$ contains a prime for all $x \geq x_0$. Schoenfeld [6] made this result precise, showing that one can take $K = 1/(4\pi)$ and $x_0 = 599$.

Cramér [1] improved the result of von Koch by furnishing the following theorem.

Theorem 1 (Cramér). *Suppose the Riemann hypothesis is true. Then it is possible to find a positive constant c such that*

$$\pi(x + c\sqrt{x}\log x) - \pi(x) > \sqrt{x} \quad (1)$$

for $x \geq 2$. Thus if p_n denotes the n th prime, we have

$$p_{n+1} - p_n = O(\sqrt{p_n} \log p_n) \quad (2)$$

Goldston [2] made this result more precise, by showing that one could take $c = 5$ in (1) for all sufficiently large values of x . He also showed that

$$p_{n+1} - p_n < 4p_n^{1/2} \log p_n$$

for all sufficiently large values of n . Ramaré and Saouter [4] improved this by showing that for all $x \geq 2$ there exists a prime in the interval $(x - \frac{8}{5}\sqrt{x}\log x, x]$.

The first purpose of this paper is to give the following improvement on the work of Ramaré and Saouter.

Theorem 2. *Suppose the Riemann hypothesis is true. Then there is a prime in the interval $(x - \frac{4}{\pi}\sqrt{x}\log x, x]$ for all $x \geq 2$.*

We prove this theorem using a weighted version of the Riemann von-Mangoldt explicit formula and some standard estimates for sums over the zeroes of the Riemann zeta-function. It will also be seen that the constant $4/\pi$ is optimal for this method and choice of weights. From the proof of Theorem 2, we go slightly further to prove the following explicit version of Cramér's Theorem.

Theorem 3. *Let $c = 2/\pi$ and suppose the Riemann hypothesis is true. Then*

$$\pi(x + (c + 1)\sqrt{x}\log x) - \pi(x - (c + 1)\sqrt{x}\log x) > \sqrt{x} \quad (3)$$

for all $x \geq 5$. Also, if p_n denotes the n th prime, we have

$$p_{n+1} - p_n < 2c\sqrt{p_n} \log p_n \quad (4)$$

for all $n \geq 1$.

Note that for ease of method, our interval takes on a slightly different form to that of Cramér, but is essentially the same as taking $c = 2 + 4/\pi$ in (1). This improves on Goldston's result.

2 Proof of Theorem 2

2.1 A smooth explicit formula

We define the von Mangoldt function as

$$\Lambda(n) = \begin{cases} \log p & : n = p^m, p \text{ is prime}, m \in \mathbb{N} \\ 0 & : \text{otherwise} \end{cases}$$

and introduce the sum $\psi(x) = \sum_{n \leq x} \Lambda(n)$. This summatory function submits itself to the Riemann von-Mangoldt explicit formula

$$\psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} - \log 2\pi - \frac{1}{2} \log(1 - x^{-2}) \quad (5)$$

where $x > 0$ is not an integer and the sum is over all nontrivial zeroes $\rho = \beta + i\gamma$ of the Riemann zeta-function $\zeta(s)$. We define the weighted sum

$$\psi_1(x) = \sum_{n \leq x} (x - n) \Lambda(n) = \int_2^x \psi(t) dt$$

and prove an analogous explicit formula.

Lemma 4. *For $x > 0$ and $x \notin \mathbb{Z}$ we have*

$$\psi_1(x) = \frac{x^2}{2} - \sum_{\rho} \frac{x^{\rho+1}}{\rho(\rho+1)} - x \log(2\pi) + \epsilon \quad (6)$$

where

$$\epsilon(x) < \frac{12}{5}.$$

Proof. We integrate both sides of (5) over the interval $(2, x)$ to get

$$\psi_1(x) = \frac{x^2}{2} - \sum_{\rho} \frac{x^{\rho+1}}{\rho(\rho+1)} - x \log(2\pi) + \epsilon$$

where

$$|\epsilon| < 2 + \left| \sum_{\rho} \frac{2^{\rho+1}}{\rho(\rho+1)} \right| + \frac{1}{2} \left| \int_2^x \log(1 - t^{-2}) dt \right|.$$

The integral can be evaluated to yield $\log(16/27)$, and the sum over the zeroes can be estimated on the Riemann hypothesis by

$$\begin{aligned}
\left| \sum_{\rho} \frac{2^{\rho+1}}{\rho(\rho+1)} \right| &< 2^{3/2} \left| \sum_{\rho} \frac{1}{\gamma^2} \right| \\
&= 2^{5/2} \left| \sum_{\gamma>0} \frac{1}{\gamma^2} \right| \\
&< 2^{5/2}(0.0233)
\end{aligned}$$

where we have used the bound (iii) from Lemma 1 of Skewes [7]. The result follows. \square

Using a linear combination of equation (6), we can probe the distribution of prime powers on an interval $(x-h, x+h)$. We do this by defining a weight function

$$w(n) = \begin{cases} 1 - |n-x|/h & : x-h < n < x+h \\ 0 & : \text{otherwise.} \end{cases}$$

One can then verify the identity

$$\sum_n \Lambda(n)w(n) = \frac{1}{h}(\psi_1(x+h) - 2\psi_1(x) + \psi_1(x-h)). \quad (7)$$

by expanding the sums on the left hand side. We insert Lemma 4 into the above equation to get the following:

Lemma 5. *Let $x > 0$ and $h > 0$ with $x \notin \mathbb{Z}$.*

$$\sum_n \Lambda(n)w(n) = h - \frac{1}{h}\Sigma + \epsilon(h)$$

where

$$\Sigma = \sum_{\rho} \frac{(x+h)^{\rho+1} - 2x^{\rho+1} + (x-h)^{\rho+1}}{\rho(\rho+1)}$$

and

$$|\epsilon(h)| < \frac{48}{5h}.$$

The error here is precisely four times the error in Lemma 4 (from the linear combination) with an inverse factor of h arising from (7). It thus remains to estimate the sum Σ . We split this sum as

$$\Sigma = \Sigma_1 + \Sigma_2$$

where Σ_1 ranges over the zeroes ρ with $|\gamma| < \alpha x/h$, where $\alpha > 0$ is to be chosen later for an optimal result, and Σ_2 is over the remaining zeroes.

For Σ_1 , we notice that its summand may be written as

$$\int_{x-h}^{x+h} (h - |x - u|) u^{\rho-1} du,$$

the absolute value of which can be bounded above by

$$\frac{1}{\sqrt{x-h}} \int_{x-h}^{x+h} (h - |x - u|) du.$$

The integral now represents the area of a triangle with a base length of $2h$ and a height of h , and so

$$\begin{aligned} \Sigma_1 &\leq \frac{h^2}{\sqrt{x-h}} \sum_{|\gamma| < \alpha x/h} 1 \\ &= \frac{2h^2}{\sqrt{x-h}} N(\alpha x/h) \end{aligned}$$

where $N(T)$ denotes the number of zeroes ρ with $0 < \beta < 1$ and $0 < \gamma < T$. By Corollary 1 of Trudgian [9], one has the bound

$$N(T) < \frac{T \log T}{2\pi}$$

for all $T > 15$ say, and so

$$|\Sigma_1| < \frac{\alpha x h}{\pi \sqrt{x-h}} \log(\alpha x/h) \tag{8}$$

so long as $\alpha x/h > 15$.

We can estimate Σ_2 trivially on the Riemann hypothesis by

$$\begin{aligned}
|\Sigma_2| &= \left| \sum_{|\gamma| \geq \alpha x/h} \frac{(x+h)^{\rho+1} - 2x^{\rho+1} + (x-h)^{\rho+1}}{\rho(\rho+1)} \right| \\
&< 4(x+h)^{3/2} \sum_{|\gamma| > \alpha x/h} \frac{1}{\gamma^2} \\
&= 8(x+h)^{3/2} \sum_{\gamma > \alpha x/h} \frac{1}{\gamma^2} \\
&< \frac{4h(x+h)^{3/2}}{\pi \alpha x} \log(\alpha x/h),
\end{aligned}$$

where the last line follows from (ii) of Lemma 1 in Skewes [7].

2.2 Primes in Short Intervals

Putting our estimates for Σ_1 and Σ_2 into Lemma 5 we have

$$\begin{aligned}
\sum_n \Lambda(n)w(n) &> h - \frac{1}{h}(|\Sigma_1| + |\Sigma_2|) - \frac{48}{5h} \\
&= h - \left(\frac{\alpha x}{\pi \sqrt{x-h}} + \frac{4(x+h)^{3/2}}{\pi \alpha x} \right) \log(\alpha x/h) - \frac{48}{5h}.
\end{aligned}$$

Notice that as we will choose $h = o(x)$, it follows that the term in front of the log is asymptotic to

$$\left(\frac{\alpha}{\pi} + \frac{4}{\pi \alpha} \right) \sqrt{x}$$

A quick hit of calculus can be used to show that $\alpha = 2$ will minimise the above term. Choosing this gives

$$\sum_n \Lambda(n)w(n) > h - \frac{2}{\pi} \left(\frac{x}{\sqrt{x-h}} + \frac{(x+h)^{3/2}}{x} \right) \log(2x/h) - \frac{48}{5h},$$

or rather

$$\begin{aligned}
\psi(x+h) - \psi(x-h) &= \sum_{x-h < n \leq x+h} \Lambda(n) \\
&> h - \frac{2}{\pi} \left(\frac{x}{\sqrt{x-h}} + \frac{(x+h)^{3/2}}{x} \right) \log(2x/h) - \frac{48}{5h}.
\end{aligned}$$

The sum on the left hand side of the above inequality is over prime powers, so we must remove the contribution of powers which are at least 2, so that only a sum over the primes remains. As such we consider the Chebyshev θ -function given by

$$\theta(x) = \sum_{p \leq x} \log p.$$

We need to bound the difference of $\psi(x)$ and $\theta(x)$ on the Riemann hypothesis. Here we can use Theorem 14 and Equation (5.5) of Schoenfeld [5] to get that

$$0.98\sqrt{x} < \psi(x) - \theta(x) < 1.11\sqrt{x} + 3x^{1/3}$$

for all $x \geq 121$. We use this bound with our inequality for $\psi(x+h) - \psi(x-h)$ to get

$$\begin{aligned} \sum_{x-h < p \leq x+h} \log p &> h - \frac{2}{\pi} \left(\frac{x}{\sqrt{x-h}} + \frac{(x+h)^{3/2}}{x} \right) \log(2x/h) \\ &\quad - 1.11\sqrt{x+h} - 3(x+h)^{1/3} + 0.98\sqrt{x-h} - \frac{48}{5h}. \end{aligned}$$

If we set $h = d\sqrt{x} \log x$, the leading term on the right hand side can be shown to be asymptotic to

$$\left(d - \frac{2}{\pi}\right) \sqrt{x} \log x + \frac{4}{\pi} \sqrt{x} \log \log x. \quad (9)$$

Thus, for $d \geq 2/\pi$ we have that there is a prime in the interval

$$(x - d\sqrt{x} \log x, x + d\sqrt{x} \log x]$$

for all sufficiently large x . We can achieve the optimal result of this method by choosing $d = 2/\pi$. Then, using a monotonicity argument one has this for all $x \geq 65000$, say. Replacing $x + d\sqrt{x} \log x$ with x , we have that there is a prime in the interval

$$(x - 2d\sqrt{x} \log x, x]$$

for all

$$x \geq 65000 + \frac{2}{\pi} \sqrt{65000} \log(65000) \approx 66798.7$$

where $2d = 4/\pi$. This completes the proof of Theorem 2, for it is easy to verify the theorem for the remaining values of x .

3 Proof of Theorem 3

The proof of this follows from the previous section. We notice that by setting $d = 1 + 2/\pi$ in (9), we have

$$\sum_{x-h < p \leq x+h} \log p \sim \sqrt{x} \log x + \frac{4}{\pi} \sqrt{x} \log \log x$$

where $h = d\sqrt{x} \log x$. Using the bound

$$\begin{aligned} \pi(x+h) - \pi(x-h) &= \sum_{x-h < p \leq x+h} 1 \\ &> \frac{1}{\log(x+h)} \sum_{x-h < p \leq x+h} \log p, \end{aligned}$$

we see that all that is required is to show that the right hand side of this inequality is greater than \sqrt{x} . The working here is straightforward and so we omit most of it. Using a monotonicity argument and the bounds from the previous section, it is straightforward to get

$$\pi(x+h) - \pi(x-h) > \sqrt{x}$$

for all $x \geq 6500$. Checking the remaining values of $x \in [2, 6500)$ takes no more than a few seconds on MATHEMATICA.

References

- [1] H. Cramér. Some theorems concerning prime numbers. *Arkiv Matematik*, 5:1–33, 1920.
- [2] D. Goldston. On a result of Littlewood concerning prime numbers. *Acta Arithmetica*, 43(1):49–51, 1983.
- [3] A. E. Ingham. *The distribution of prime numbers*. Number 30. Cambridge University Press, 1932.
- [4] O. Ramaré and Y. Saouter. Short effective intervals containing primes. *Journal of Number Theory*, 98(1):10–33, 2003.
- [5] J. B. Rosser and L. Schoenfeld. Approximate formulas for some functions of prime numbers. *Illinois Journal of Mathematics*, 6(1):64–94, 03 1962.
- [6] L. Schoenfeld. Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$. II. *Mathematics of Computation*, pages 337–360, 1976.
- [7] S. Skewes. On the difference $\pi(x) - \text{li}(x)$ (II). *Proceedings of the London Mathematical Society*, 3(1):48–70, 1955.
- [8] E.C. Titchmarsh. *The Theory of the Riemann Zeta-function*. Oxford University Press, second edition, 1986.
- [9] T. S. Trudgian. An improved upper bound for the argument of the Riemann zeta-function on the critical line. *Mathematics of Computation*, 81(278):1053–1061, 2012.
- [10] H. von Koch. Sur la distribution des nombres premiers. *Acta Mathematica*, 24:159, 1901.